THE DEVELOPMENT OF A TEST FOR SELECTING RESEARCH PERSONNEL

The second in a series of reports prepared under the sponsorship of the

Manpower Branch
Human Resources Division
Office of Naval Research

January 1950

AMERICAN INSTITUTE FOR RESEARCH

Pittsburgh, Pennsylvania

41

THE DEVELOPMENT OF A TEST FOR SELECTING RESEARCH PERSONNEL

The second in a series of reports prepared under the sponsorship of the

Manpower Branch

Human Resources Division

Office of Naval Research

Project Director

Mary H. Weislogel

Project Staff

Ernest R. Kolovos
Marilyn K. Rigby

Advisory Staff

John C. Flanagan, Director of Research
Geraldine Spaulding, Test Editor
Elmer D. West Ralph F. Wagner

January 1950

AMERICAN INSTITUTE FOR RESEARCH

C O N T E N T S

Technical Appendices

Technical appendices for this report are published separately. They contain material primarily of technical interest. The appendices include project memoranda such as instructions for item-writers and instructions for test critics; a list of the critical requirements for research personnel; rationales for the preparation of test items to predict the critical behaviors; and item-analysis data obtained from trial administrations of test items.

A C K N O W L E D G M E N T S

The work described in this report was carried out under contract with the Office of Naval Research. Throughout the project, the staff worked closely with the Manpower Branch of the Human Resources Division. Dr. Ralph M. Hogan, Head of the Manpower Branch, provided valuable assistance and advice throughout the project. The later phases of the project were carried out under the general cognizance of Dr. J. W. Macmillan as the head of the Human Resources Division.

Special acknowledgments are due to those who undertook the difficult task of item-writing. Dr. Meredith P. Crawford, Dr. Palmer O. Johnson, Dr. Truman L. Kelley, Dr. Irving Lorge, Dr. J. William Maucker, Dr. Walter R. Miles, Dr. Robert B. Miller, Dr. Charles I. Mosier, Marion F. Shaycoft, Dr. Lawrence M. Stolurow, Dr. Robert L. Thorndike, Dr. Robert M. W. Travers, and Dr. Joseph Weitz each prepared a number of items. David Boodman, Clifford Swensen, and Alfred Weber also prepared a smaller number of items. In addition to the work of the regular project staff in item-development, Dorothy L. Berger served as a member of the project staff for a time and prepared a number of items.

Dr. Truman L. Kelley and Charlotte C. Davis provided valuable editorial assistance during the early phases of the project and Dr. Richard E. Watson went over the tryout form of the test as an editorial critic. Dr. Frederick B. Davis, Dr. Paul Horst, and Dr. Dael Wolfle assisted by reviewing the items, including the rationales of the item-development process.

The cooperation of personnel in the Navy laboratories in reviewing the technical content of test items is gratefully acknowledged. Dr. W. G. Torpey, Acting Personnel Officer of the Naval Research Laboratory, and William Oncken, Jr., Training Director of the Naval Ordnance Laboratory, provided valuable suggestions and made arrangements for interviews with laboratory personnel serving as critics. The assistance of Dr. Oswald Blackwood, Dr. G. R. Fitterer, Dr. William E. Wallace, and James D. Minford of the University of Pittsburgh in reviewing the subject-matter content of test items was also highly valuable during editing phases of the project. Special thanks are due to Louis J. Sparvero and the Pittsburgh Naval Research Reserve Unit (4-3) for assistance in the review of test items. Milton M. Mandell of the Test Development Division, United States Civil Service Commission, furnished valuable suggestions and advice.

Considerable time and personal effort were devoted in arranging for administration of test items by Dr. William A. Owens of Iowa State College; Dr. Herbert Longenecker, Dr. G. R. Fitterer, Dr. Ralph Jensen, and Dr. James S. Taylor of the University of Pittsburgh; and Roland C. Moore and Dr. Joseph Weitz of the Carnegie

Institute of Technology.

The project was carried out under the direction of Dr. John C. Flanagan. Dr. Elmer D. West, Geraldine Spaulding, and Dr. Ralph F. Wagner assisted in planning throughout the project. Editing of test items was supervised by Geraldine Spaulding. Item editing, test administration, and data analysis, as well as preparation of test items, were the functions of the project staff, Ernest R. Kolovos, Marilyn K. Rigby, and Mary H. Weislogel.

This report was written by Mary H. Weislogel. The diagrammatic and graphic materials were prepared by Ruth V. McBride; tabular materials were prepared by Jane C. Little.

SUMMARY

The present report concerns a test development project carried out as part of a broad research program concerning scientific personnel, sponsored by the Manpower Branch, Human Resources Division, of the Office of Naval Research. The long-range program would include:

1. Determination of the critical requirements for successful participation in research and engineering work,
2. Development of an aptitude test for the selection of scientific personnel,
3. Development of tests to measure proficiency in specific areas of scientific work,
4. Development of procedures for evaluating the job performances of scientific personnel.
5. Determination of the predictive value of the tests developed in steps (2) and (3) using the procedures developed in step (4) to obtain evaluations of personnel for comparison with test predictions.

The first of these steps was completed in an earlier project. The second step, the development of an aptitude test for the selection of scientific personnel, is the subject of the present report. The objectives of the project were, first, that the test developed be applicable to the selection of candidates for advanced training in natural science and engineering, as well as to the selection of junior professional workers in research laboratories; second, that it be a test of potentiality for, rather than proficiency in, research work; and third, that each test item developed be an attempt to predict a specific critical behavior identified in a preceding study concerning the critical requirements for research personnel.

The project resulted in the development of a test of approximately 170 items, organized in three groups or subtests; subtest I is intended to predict chiefly job performances related to formulating problems and hypotheses and planning and designing the investigation; subtest II is intended to predict behaviors related to conducting the investigation and interpreting research results; subtest III, behaviors related to preparing reports, administering research projects, and accepting organizational and personal responsibility. Since security of the actual content of test items must be maintained, items contained in the final form of the test are not given here. However, sample items illustrative of those contained in the test are presented.

One of the principal contributions of this study has been the development of detailed hypotheses or "rationales" regarding the most suitable approach to the measurement of each of the critical requirements for research personnel established

in the previous study. The appropriateness of these items as predictors of their corresponding critical behaviors has been reviewed by both experts in this type of research and by professional consultants in the field of test-construction. Thirty-six of the rationales developed for writing items are presented in this report. These include samples from each of the thirty-six basic categories of the critical requirements in terms of behavior.

It is believed that the test will be of value in selecting research personnel for the following reasons:

1. Test items were developed to predict specific behaviors identified by scientists themselves in a preceding study as crucial to successful performance in research work.
2. Items were prepared by individuals with knowledge of both scientific subject-matter and the technical aspects of test-construction.
3. Items were criticized by persons professionally employed in research work in natural science and engineering, and were revised on the basis of such review.
4. Trial administrations were conducted to determine the difficulty of items and to identify further editorial needs.
5. Items were reviewed by editorial consultants with particular reference to whether each would predict the intended critical behavior.

The test of research aptitude developed in this project has been reproduced and is available for use in the later phases of the program outlined above.

CHAPTER I

THE PROBLEM

Introduction

The experience of our government and military services during World War II led to an unprecedented development of research in the natural sciences and engineering. Directly related to this growth was a tremendous increase in research and development programs directly supervised by the government, and a consequent increase in the numbers of scientific personnel employed by the government. In the postwar period, the activities of government technical laboratories have retained their importance; it is considered essential, both for purposes of national defense and peacetime development, that major efforts in research and engineering be continued.

Rapid expansion of this program has naturally brought new problems to administrators of government research laboratories, and particularly to the personnel departments. They have been faced with the task of securing large numbers of qualified personnel; with placing personnel in positions appropriate to the individual's education and experience; with providing supplementary and on-the-job training for specific assignments; and with evaluating the effectiveness of personnel on the job.

Administrators in the laboratories and governmental groups concerned with the formulation and coordination of policies affecting scientific personnel have recognized the need for research and development in personnel procedures. These groups have pioneered in planning and initiating needed personnel research. Among these, the Office of Naval Research, through its Human Resources Division, has been active in formulating and subsidizing a personnel research program. The present report concerns a development project carried out as a part of this program.

Probably the most basic problem in the effective utilization of scientific personnel is selection. There are two aspects of the selection problem: first, to identify those individuals who have the potentialities both for making a contribution to some aspect of the over-all technical program and for developing their capacities so as to grow into more responsible positions; second, to determine whether the individual possesses the necessary technical knowledge to fill a specific position for which he is being considered. The first requires a "screening" device which will identify individuals of high caliber who are capable of developing into candidates for advanced positions in research and engineering work. Such a device could be used to sort applicants into groups according to the level of performance to be expected from each. The most economical device would be one

which could be applied equally to individuals with training in varied subject-matter specialties. It would also be economical to apply such a measure before advanced training is completed; if selection were made at the college senior level, those candidates for advanced training who would profit most from an educational subsidy could be identified.

The second problem requires tests of knowledge or proficiency in the field in which employment is sought. In contrast to the first, these tests would be directed towards technical background in a specific field of specialization. They would also differ from the first in measuring achievement of specific knowledges or skills rather than potentiality for achievement.

The Long-Range Research Program

A long-range research program has been planned in cooperation with the Office of Naval Research to study these basic personnel problems. The logical first step was considered to be a comprehensive study to determine the job behaviors which are crucial for successful participation in research and engineering work. This study would furnish the basic data to which all the other steps in the research program would be referred. The second step would be the development of an aptitude test to be used in selecting scientific personnel, and the third step, the development of tests designed to measure technical proficiency on standard job samples. Such tests would be applicable to the selection and placement of scientific personnel. A fourth step would be the development of procedures for evaluating the job performances of personnel. Such evaluation would be used as the criterion of job efficiency in determining the predictive value of the aptitude and proficiency tests developed. Finally, the fifth step would involve administration of the aptitude and proficiency measures to large groups of prospective scientific personnel and following up on their later success as judged by the procedures for evaluation of job performance developed in step four. The initial work on the fifth step would be done concurrently with the third and fourth steps.

The first of these steps, the determination of the crucial job behaviors or "critical requirements" for research personnel, has been completed in an earlier project.[1] The basic data in that study were job behaviors observed by scientists and judged by them as crucial to success or failure on the job. Three thousand descriptions of actual situations occurring in research installations were obtained. These specific incidents described by scientific workers themselves were then classified and formulated into a list of critical requirements. The list contains behaviors associated with a variety of jobs in research and engineering work --in basic and applied research, design, development, testing, and evaluation;

---

[1] Critical Requirements for Research Personnel. The American Institute for Research, Pittsburgh, Pa., 1949.

jobs in most of the natural sciences and engineering are included. A total of 428 different behaviors were identified and organized in groupings of eight major areas[2] and thirty-six subareas.

Objectives of the Present Project

The second step in this long-range research and development program was next undertaken, and is the subject of the present report. The basic objective of the project was the development of a test of potentiality or aptitude applicable to the selection of research personnel for the natural sciences and engineering. It was believed that this objective could be reached most effectively by developing test items directly related to the critical behaviors identified in the preceding study. The scope of the project was considered to extend only as far as the development, preliminary tryout, and revision of the test. Determination of its predictive value would be the objective of a later project.

As a preliminary to test planning, it was first necessary to define the project objective in specific terms. It was believed that selection of research personnel would be most effectively accomplished if selection were made at the college senior level. The major purpose of the test was therefore specifically defined as the selection of candidates for advanced training in natural science and engineering from the college senior population. However, it was expected that the items would also be applicable to the selection of junior professional personnel in research and engineering laboratories.

It was intended that the test measure aptitude or potentiality for success in research work rather than the possession of specific knowledge or technical proficiency. It was recognized that tests of subject-matter knowledge would be useful in predicting success, but the objective here was to identify individuals with potentiality for successful performance in research or engineering regardless of the field of specialization. Such an approach would permit selection at an early stage of professional training. It would also be useful for a variety of subject-matter specialties, whereas proficiency measures would necessarily be concerned with a particular field.

The specific objectives of the test development project, then, were: first, that a test applicable to the selection of candidates for advanced training in natural science and engineering from the college senior population be developed; second, that it be a test of potentiality for, rather than proficiency in, research work; and third, that each test item developed be an attempt to predict a specific critical behavior identified in the preceding critical requirements study.

---

[2] The eight major areas are: I Formulating Problems and Hypotheses; II Planning and Designing the Investigation; III Conducting the Investigation; IV Interpreting Research Results; V Preparing Reports; VI Administering Research Projects; VII Accepting Organizational Responsibility; VIII Accepting Personal Responsibility.

CHAPTER II

P L A N N I N G O F T H E T E S T

Working Assumptions

Before the work of test planning and item construction proper was undertaken, specific working assumptions were formulated. Several of these assumptions have already been implied in the statement of project objectives. The first and most basic assumption was that the critical behaviors identified in the preceding study represent the most important job behaviors of scientific personnel and should therefore serve as the basis for building test items. This assumption follows directly from the methodology used in the identification of the critical behaviors. Since a large sample of senior research and engineering personnel reported observations of job behaviors identified by them as crucial to success or failure, the resulting list of 428 critical behaviors is believed to cover the important aspects of scientific work.

A second, related assumption was that most of these job behaviors are important for all the subject-matter areas found in research and engineering work. Data obtained in the preceding study indicated that the various critical behavior categories were significant in many fields of specialization. That is, there was little tendency for the critical behaviors for one area to be concentrated in any one subject-matter field. Since most of the critical behaviors are important for the different scientific fields, it was further assumed that aptitudes underlying the critical behaviors would also be important for all the scientific fields.

A third assumption involved the place of the critical behaviors in the rationale for the construction of an aptitude test. It was felt that items should be constructed to predict specific critical behaviors rather than to measure more general aptitudes inferred to underlie a group of behaviors. That is, each test item would grow directly from a critical behavior by posing a problem situation requiring that particular behavior to be demonstrated.

A fourth assumption was necessary to insure that items would in fact measure potentiality rather than subject-matter proficiency. Items must be couched in terms of a specific subject-matter, and those students specializing in the particular field might have an advantage through greater familiarity with specialized terminology, techniques, and principles. However, if each examinee had equal opportunity to acquire technical information pertinent to the problem, such special advantages would be lessened. In an effort to insure such equality of opportunity, it was considered necessary to specify the content and level of training which all

examinees would have in common. It was assumed that there would be a minimum common background of one year's college training in physics, chemistry, and mathematics for all students in most areas of natural science or engineering. Any necessary technical information beyond this level would be supplied in the statement of each problem.

This procedure would operate to minimize the advantage of the examinee trained in the subject-matter field covered by a specific item. In addition, the total test would be balanced in technical content, so that no one field of specialization would be involved in a majority of items. Most of the items would be couched in terms of physics, chemistry, mathematics, or engineering subject-matter. A limited number might involve situations in biological science, since natural science and engineering students could be expected to have approximately equal familiarity with this area.

A final assumption was made regarding the type of test item considered most applicable. The test was intended to be of the objective, machine-scorable type suitable for administration to large groups. The multiple-choice item with five alternatives was considered most suitable for the project purposes. However, all varieties of objective test items were considered throughout the item-development phases.

Survey of the Literature and the Test Outline

As a preliminary to further test planning and construction of items, a study of current literature pertinent to the problem was undertaken. The reader is referred to the bibliography at the end of this report for a list of references consulted.

The first part of this phase consisted of a brief survey of existing objective tests designed to measure some aspect of scientific aptitude. The purpose of examining these tests was to review the various types of items that had been previously developed. Abstracts concerning ideas suggested by these tests were prepared, containing a general description of the type of information that might be presented in the statement of the problem, type of question to be posed, and types of alternatives from which the right answer would be selected.

Each critical behavior was also studied, and ideas as to appropriate types of test items for each were formulated. These notes were combined with the file obtained through study of pertinent tests, resulting in a set of file cards containing ideas for test items, coded to specific critical behaviors. An illustration is given in Figure 1 on page 6.

The final step in the planning process involved preparation of a test outline indicating the number of items to be prepared for each category of critical behaviors. Data obtained from the critical requirements study which preceded this

Figure 1

SAMPLE FROM FILE OF ITEM-TYPES

Major Area:	II. Planning and Designing the Investigation
Subarea:	C. Identifying and Controlling Important Variables
Critical Behavior:	1. OUTLINED PLAN PERMITTING CONTROL AND SYSTEMATIC VARIATION OF ALL RELEVANT VARIABLES.

SUGGESTIONS FOR ITEM TO PREDICT THIS CRITICAL BEHAVIOR:

Information to be given: explanation of problem and its variables; description of design for an experiment in which an important variable is not properly controlled.

Problem to be set for examinee: select the change in the design of the experiment that would result in the greatest improvement in design.

Alternative choices to be given: a list of five possible changes in the experimental design, the correct choice being a more rigorous control of the variable that is not properly controlled.

project were directly applicable to the test outline. The critical behaviors were organized under eight major areas and thirty-six subareas. The test outline, indicating number of items needed for specific categories, was broken down in terms of these large categories rather than the smaller critical behavior groupings. The relative proportion of critical behaviors reported for each major area in the preceding study was used as a rough guide to the number of items required. The numbers of critical behaviors reported for each subarea were also available for reference. The percent of the total number of critical behaviors obtained for each major area and its subareas is shown in Table I.

Table I

PERCENT OF CRITICAL BEHAVIORS IN EACH CATEGORY

Note: Area I has only three subareas; Area III has only six; Area IV has only two, etc.

	Area	Subareas						
Area	Total	A	B	C	D	E	F	G
I	5.0	1.8	1.7	1.5				
II	17.3	4.6	1.1	3.4	3.2	3.0	1.4	0.6
III	22.6	8.9	4.1	1.3	2.9	5.0	0.3	
IV	7.2	5.3	1.9					
V	16.1	5.6	7.4	1.9	1.0			
VI	11.3	2.4	2.9	2.8	1.8	1.4		
VII	9.7	4.5	1.9	1.8	1.6			
VIII	10.8	0.9	3.9	0.2	2.9	2.8		

CHAPTER III

DEVELOPMENT OF THE TEST

Item-Writing

Individuals with knowledge of both scientific subject-matter and the technical aspects of test-construction were selected to prepare items. A number of field personnel as well as project staff participated in the construction of items. Item-writers were given detailed information and instructions regarding the types of items desired. Each received a list of the critical behaviors and a memorandum explaining the background and rationale of the project, and containing specific instructions.[1] In order to assure a balanced coverage of all the behavior areas, each item-writer was assigned a specific major group of critical behaviors for which to prepare items.

The general procedure suggested to item-writers was that they first study the critical behavior and then prepare an item for it based on an original situation. Source materials might be consulted when necessary, but would be altered so that examinees familiar with the original source would not have a special advantage.

Preliminary Editing

Preliminary editing of items was carried out by a senior editor and the project staff with the assistance of two editorial consultants. During the editing process, the following major criteria were used in evaluating items:

1. Is it probable that the item will predict success or failure on the job as defined by the critical behaviors?
2. Is the difficulty-level of the item appropriate to the group for which the test is intended?
3. Is the answer intended as correct clearly the best choice of those offered?
4. Does solution or the item require only such academic information as is ordinarily acquired in one-year college courses in physics, chemistry, and mathematics?
5. Does the item meet the generally accepted technical requirements of good test items?

With these criteria in mind, the editorial staff first independently reviewed each item, noting comments, questions, or suggested changes. These editorial comments were then discussed and decisions were made concerning preliminary changes. Even with these items prepared by highly competent individuals, considerable edi-

[1] This memorandum is contained in the Technical Appendices for this report.

torial change was found to be necessary. To illustrate the editing process, a preliminary item and editorial comments are shown in Figure 2 below. The item is given with changes indicated, and the editorial comments which were the basis for these changes are shown below the item.

Figure 2

SAMPLE OF PRELIMINARY EDITING

ITEM:

You have recently been appointed Director of ~~a~~ *an industrial* research laboratory. At the present time it consists of eight divisions with from four to ten professional workers in each division. It is ~~the Director's~~ *your* responsibility to coordinate the work of these divisions, to see that they work together, and to see that each makes an effective contribution to the work of the laboratory as a whole. Which of the following procedures should you use to obtain the most adequate coordination and supervision?

A. Appoint ~~a~~ supervisor ~~from~~ each division *the man with the best record in* who will report directly to you at regular intervals.

B. Select from the laboratory as a whole several men ~~with excellent~~ records and appoint them as supervisors directly responsible to you. *who have the best*

C. Create a separate supervisory division whose members will work in all divisions at different times and will be directly responsible to you.

X D. Designate an administrative supervisor in each division, and arrange to meet regularly with these men as a group. ~~to consider general problems of coordination and supervision.~~

E. ~~Bring in as head of each division a scientist from the outside, who will feel a personal loyalty and a direct responsibility to you.~~ *Reorganize the laboratory into twelve divisions so that individual professional workers will receive more adequate supervision.*

EDITORIAL COMMENTS:

Text of item:	Type of laboratory should be specified.
	Change "the Director's responsibility" to "your responsibility" to give impression examinee is actually in the situation.
Choice A:	Could be made more attractive by specifying that individual with the best record is to be appointed.
Choice D:	Correct choice, being longer than others, might stand out.
Choice E:	The idea of "bringing in an outsider" may make it unattractive; suggest "reorganize the laboratory into twelve divisions, so that workers will receive more adequate supervision."

Items with the technical content and level of difficulty necessary for a test of this type are very difficult to prepare, and many items had to be discarded. Approximately 50 percent of the items written were retained in the final form.

Subject-Matter Review

When items had undergone preliminary editing, they were next reviewed by persons professionally engaged in some area of natural science or engineering. The major purpose of this review was to insure that all technical subject-matter contained in the items was properly used and within the first-year-college level of information assumed for the examinees. In addition, critics were asked specifically to consider the following two questions:

1. Will solution of the test problem predict the critical behavior for which it was developed?
2. Is the answer intended as correct the only correct choice (or clearly the best choice) of those offered?

Items to be used in preliminary tryouts were reviewed by at least three subject-matter critics, in order that pooled judgments might be obtained. Scientific personnel from the Naval Research Laboratory, Naval Ordnance Laboratory, Naval Research Reserve Unit (4-3), and faculty of the Departments of Chemistry, Physics, and Metallurgical Engineering at the University of Pittsburgh served as critics.

Comments of these critics were the basis for additional editorial revision. An illustration of the kinds of comments obtained is given in Figure 3 below. The item criticized is shown on the left, and the comments of subject-matter critics concerning that item are shown on the right.

Figure 3

SAMPLE OF CRITICS' COMMENTS

Major Area: II. Planning and Designing the Investigation
Subarea: E. Developing Plans for the Use of Equipment, Materials, or Techniques
Critical Behavior: 1. INCLUDED IN PLAN OF INVESTIGATION EQUIPMENT, MATERIAL OR TECHNIQUES WHICH MET THE REQUIREMENTS OF THE PROBLEM.

ITEM	CRITICAL COMMENTS
It is desired to develop an incendiary bomb that will contain, among other things, a material that will increase the violence of conflagration when water is sprayed or poured upon it. This material should have the following characteristics: it should be light in weight, easily melted for shell-filling, fairly plentiful, and it should not, itself, be a mixture. Which of the following materials would you recommend? A. Phosphorus B. Sulphur C. Magnesium X D. Potassium E. Oxygen	1. A chemist would have a greater advantage in this item than physicists, mathematicians, or engineers. 2. An individual with only one year of work in chemistry could not be expected to remember this information. 3. Individuals with specific work experience (for example, war experience) would solve the item more easily.

(This item was discarded because the editorial staff felt it could not be revised sufficiently to correct the indicated difficulties.)

Trial Administrations

Following completion of preliminary editing and review of items by subject-matter critics, trial administrations of groups of items were conducted. The purpose of these tryouts was twofold. The first objective was to determine the difficulty of each item for the type of group for which the test was intended; this information would be useful in arriving at a final set of items constituting a test with a difficulty-level appropriate for selecting candidates for advanced training in natural science and engineering. The second purpose was to identify need for further editorial work with specific items.

A number of different groups of students at three colleges and universities participated in the trial testing sessions. During the earlier tryouts, several short forms of the test were used. The content and length of test, composition of the group, and available testing time varied considerably during these sessions. The following groups of students took early forms of the test: 67 seniors and graduate students in natural science and engineering at the Carnegie Institute of Technology; 8 graduate students in natural science at the University of Pittsburgh; and 13 seniors and 17 graduate students in Metallurgical Engineering at the University of Pittsburgh. Data obtained from these early tryouts were used as the basis for further editorial revision in the preparation of final tryout forms.

The final tryout forms were administered to 189 students in mathematics, natural science, and engineering at the University of Pittsburgh - largely seniors and graduate students - and 130 students in mathematics, natural science, and engineering at Iowa State - also largely seniors and graduate students. The forms used for these groups also differed, although the item content of the forms overlapped. Two forms of about 40 items each were used for the Pittsburgh group, half of the group taking one form and half taking the other form. Two different 90-item forms were used for the Iowa group with half of the group taking one form and half the other. The time allotted for the Pittsburgh group was 50 minutes; the Iowa group was allowed a maximum of three hours.

Examinees were instructed to answer items in the order given, and to answer questions about which they were not absolutely sure, but to avoid sheer guessing. The raw score for each individual was therefore corrected by subtracting one-fourth of the wrong answers from the number of right answers. The distributions of scores are shown separately for the two groups in Figure 4 and Figure 5. Scores of the two groups are not comparable because of the differences in test content, maximum possible score, and length of test session described above.

Data showing the distribution of the items according to difficulty-level were also tabulated. The difficulty index represents the percent of examinees who chose the correct response for the item, and is calculated simply by dividing the number of examinees selecting the correct response by the total number answering the item.

Figure 4

SCORES FROM UNIVERSITY OF PITTSBURGH FINAL TRYOUT

(Students in mathematics, natural science, and engineering, largely seniors and graduate students)

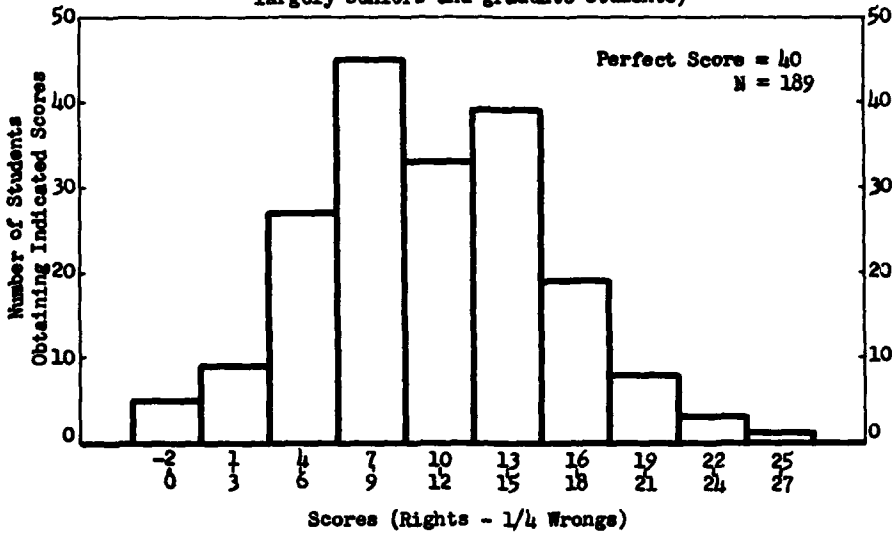
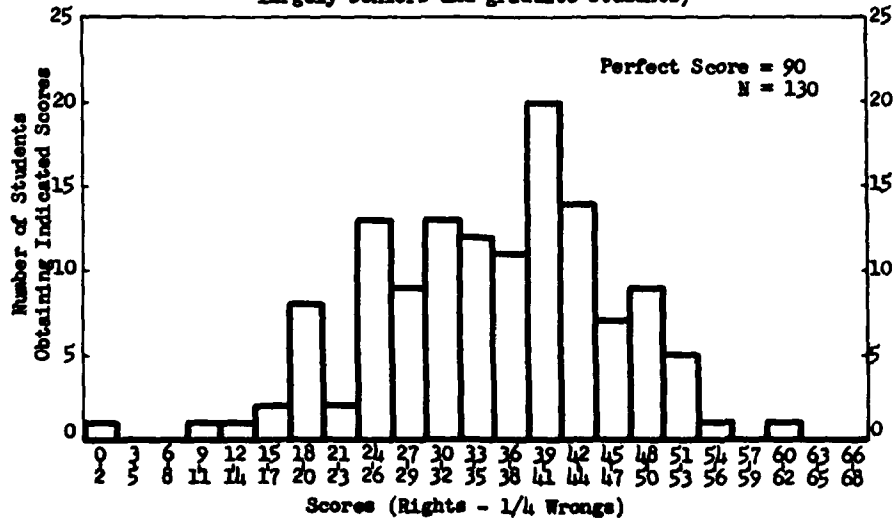


Figure 5

SCORES FROM IOWA STATE COLLEGE TRYOUT

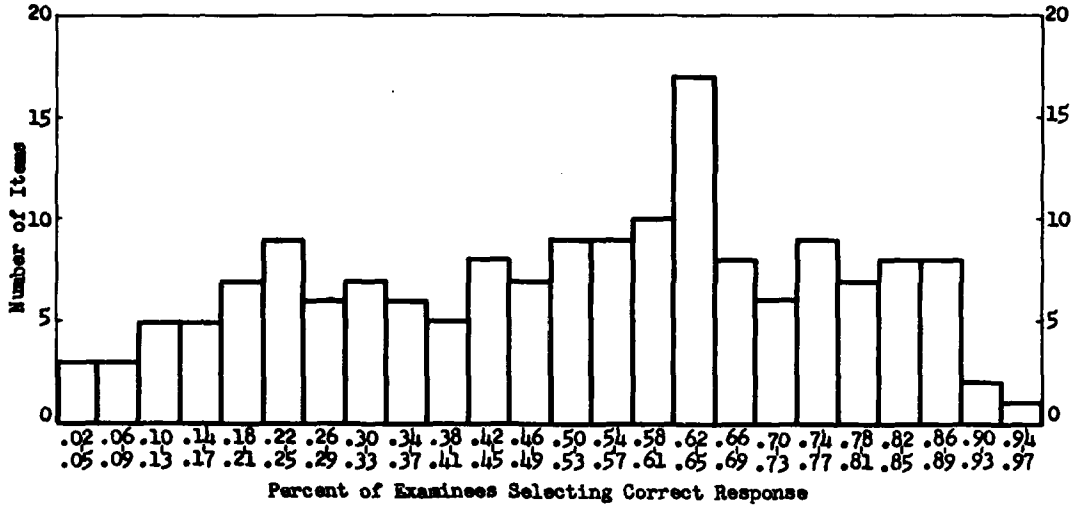
(Students in mathematics, natural science, and engineering, largely seniors and graduate students)



Data obtained from the earlier tryouts were included with the data from final tryouts whenever the particular item remained unchanged in both earlier and final tryout forms. The difficulty-levels are shown in Figure 6, which indicates the number of items found in each class of difficulty.

Figure 6

DIFFICULTY-LEVELS OF ITEMS



Results of the trial administrations proved very useful in further revision of test items. A number of items had difficulty indices which suggested that most examinees could be expected to respond correctly. Such items would have little predictive value, and were therefore revised whenever possible.

The results of the trial administrations furnished an additional type of data helpful in item editing. The number of examinees selecting each response to an item was tabulated,[2] and analysis of these data offered clues to further editorial needs. For example, if analysis showed that responses were about equally divided between the correct response and one of the wrong alternatives, the item would require re-examination to determine whether the wrong alternative was too attractive or possibly an acceptable response. If either of these suppositions were likely, changes would be necessary in order to correct the difficulty.

[2]Tabulations of these data are contained in the Technical Appendices for this report.

Final Editing and Critical Review

Each item was reviewed by the editorial staff in the light of data obtained from the trial administrations. In addition, the final tryout forms were reviewed independently by four editorial consultants. One of these critics was particularly qualified to criticize items regarding their subject-matter content. The other three were especially well qualified to review the test from the point of view of the technical requirements of good test items. Final editing was completed by the staff on the basis of these critics' comments and data obtained from the final tryouts.

CHAPTER IV

RATIONALES AND SAMPLES OF TEST ITEMS

The final form of the test consists of approximately 170 items, which are organized in three groups or subtests. These subtests contain items designed to predict critical behaviors in the following areas: Subtest I, Formulating Problems and Hypotheses and Planning and Designing the Investigation; Subtest II, Conducting the Investigation and Interpreting Research Results; Subtest III, Preparing Reports, Administering Research Projects, and Accepting Organizational and Personal Responsibility.

It was considered desirable to summarize the rationale of the test in concrete terms by indicating how each critical behavior had been analyzed and an item developed to predict that behavior. The reader will recall that one of the first steps undertaken was the preparation of a file of item-types based on study of each critical behavior. Item-writers had also been asked to indicate why they believed each item would predict the intended critical behavior.

With the help of these sources, specific statements concerning the meaning of the critical behaviors and the logical relationships between test items and the behaviors were developed. These statements or "rationales" took the following form: first, a formulation of hypotheses as to the knowledges, abilities, or personality traits believed to underlie each critical behavior; second, a description of the type of test item which would appropriately sample these knowledges, abilities, or traits; and third, an analysis of the item actually developed for each specific critical behavior; this analysis was intended to show how the item fulfills the requirements postulated in the first two parts of the rationale.

The rationales were reviewed and criticized by the three editorial consultants mentioned above in conjunction with their review of test items. Samples of the rationales are shown on the pages which follow. They have been selected so that each of the thirty-six subareas is represented.[1] Sample test items, one for each of the eight major areas, are included following a few of the rationales. Since security of items must be maintained, the texts of those contained in the test proper are not given in this report. The sample items, however, are illustrative of the types contained in the test.

The rationales and sample test items which follow are arranged in the order of the major areas, subareas, and critical behaviors[2] of the list of critical re-

[1] The complete series of rationales is contained in the Technical Appendices for this report.

[2] The list of critical behaviors is contained in the Technical Appendices.

quirements obtained from the preceding study. Each rationale contains: (1) a statement of the major area (Roman numeral), subarea (capital letter), and critical behavior (arabic numeral) concerned; (2) a statement of hypotheses as to the knowledges, abilities, or personality traits believed to underlie that critical behavior, under the heading, "This behavior seems to involve"; (3) a description of the type of test item it is believed would appropriately sample these knowledges, abilities, or traits, under the heading, "An item intended to predict this behavior." In those cases where a sample item is presented, this description relates specifically to the item, and is intended to show how the item fulfills the requirements postulated for prediction of the critical behavior; the sample item is then presented directly below, with the intended answer given following the item.

Rationales and Sample Items

I. Formulating Problems and Hypotheses
   A. Identifying and Exploring Problems
      4. SUGGESTED A NEW PROBLEM WHICH COULD BE STUDIED WITH AN ALREADY SUCCESSFUL TECHNIQUE.

This behavior seems to involve:

(1) Familiarity with a technique that has already been used successfully
(2) The ability to abstract a class of problems whose requirements would be fulfilled by that technique
(3) The ability to select one problem from that class

An item intended to predict this behavior:

The item below describes a technique that has been used successfully and then lists several problems. The examinee is asked to select the problem for which the technique probably has most value. This should sample (2) and (3) above.

Artificially produced radioactive isotopes have given the scientist a new tool of great value in certain types of investigation. Tracer techniques involve putting artificially produced radioactive isotopes of known identity and chemical form into the system under study. Later the presence and chemical form of the isotope is checked for at various places in the system and from the information obtained, inferences as to the intervening mechanisms are made. The unique advantage of these isotopes in this application lies in the fact that they possess the same chemical characteristics as isotopes usually present, but retain their radioactive "tags" throughout all chemical transformations.

In which of the following problems would isotopic tracer techniques probably be of most value?

A. The migration of microorganisms in natural bodies of water
B. The movement of air masses over the North American continent
C. The age or period of growth of trees
D. The heat exchange phenomena in internal combustion engines
E. The composition of stars

(Intended Answer: A)

I. Formulating Problems and Hypotheses
B. Defining the Problem
3. GATHERED INFORMATION ON EXACT REQUIREMENTS, SPECIFICATIONS, AND GOAL OF ASSIGNED PROJECT.

This behavior seems to involve:

(1) Recognizing the importance of gathering information
(2) Identifying information required by problem
(3) Taking steps to gather information
(4) Taking these steps in proper order

An item intended to predict this behavior:

Describe a problem, present five elements of information that solution of problem may require, and ask examinee to select the one that would logically be sought first. This would test (4) above.

- - - - -

I. Formulating Problems and Hypotheses
C. Setting Up Hypotheses
1. PROPOSED HYPOTHESIS OR GENERAL FORMULA IN ORDER TO EXPLAIN OBSERVED PHENOMENA.

This behavior seems to involve:

(1) The ability to think of potential explanations of an observed event
(2) The ability to evaluate the explanations
(3) The ability to put the explanations in order of promise or practicability

An item intended to predict this behavior:

Present a description of an observed event and list five explanatory hypotheses, one of which is considerably less probable than the others, and require the examinee to select the least promising hypothesis; or present one promising and four poor hypotheses, requiring that the best be selected. Such an item would require the examinee to perform steps (2) and (3) above.

- - - - -

II. Planning and Designing the Investigation
A. Collecting Background Information
1. SOUGHT OUT INFORMATION AND IDEAS FROM EXISTING LITERATURE, ASSOCIATES, OR EXPERTS ON PROBLEM BEFORE BEGINNING WORK ON PROJECT.

This behavior seems to involve:

(1) Understanding that literature, associates, and experts may supply information that allows unnecessary work and errors to be avoided
(2) Understanding that the sooner these sources are consulted, the less unnecessary work will be done and the fewer errors committed

An item intended to predict this behavior:

Describe a specific technical task to be performed. Indicate that the person who is to perform the task is unfamiliar with it. As alternatives offer various ways to begin work on the problem. Examinee is required to select consulting literature as preferable. This should test (1) and (2).

- - - - -

II. Planning and Designing the Investigation
B. SETTING UP ASSUMPTIONS

This behavior seems to involve:

(1) The ability to formulate assumptions appropriate to the problem
(2) The ability to examine assumptions critically
(3) The ability to select assumptions appropriate to the problem on the basis of (2)

An item intended to predict this behavior:

Describe an experimental design and list five possible assumptions. Ask the examinee to select the assumption in the list that is necessary to the design. This should test (2) and (3).

- - - - -

II. Planning and Designing the Investigation
C. Identifying and Controlling Important Variables
1. OUTLINED PLAN PERMITTING CONTROL AND SYSTEMATIC VARIATION OF ALL RELEVANT VARIABLES.

This behavior seems to involve:

(1) The ability to recognize relevant variables
(2) The ability to decide which variables must be controlled in order to test a hypothesis adequately
(3) The ability to devise methods to control the variables selected in (2)
(4) The ability to integrate the materials in the development of a plan

An item intended to predict this behavior:

Describe an experiment where all relevant variables are not properly controlled. Present five alternatives offering different experimental designs, the correct choice being the one which best controls relevant variables. Ask the examinee to select the best plan; in doing so he must, in part, perform steps (1), (2), and (3).

- - - - -

II. Planning and Designing the Investigation
D. Developing Systematic and Inclusive Plans
5. OUTLINED PLAN CALLING FOR EACH ELEMENT OF PROBLEM TO BE STUDIED IN SEQUENCE.

This behavior seems to involve:

(1) Knowing or developing a definition of the problem
(2) Selecting or developing a method appropriate to the problem
(3) Outlining in detail steps involved in the method
(4) Selecting the appropriate first step

An item intended to predict this behavior:

Describe a problem and offer as alternatives, steps that might be taken in various procedures. The examinee must select the step that would come <u>first</u> in the <u>most</u> appropriate method. This should test (2) and (4).

- - - - -

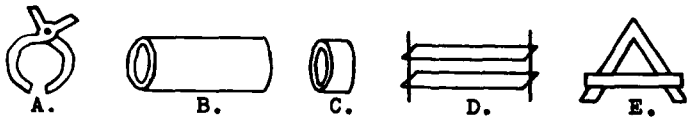
II. Planning and Designing the Investigation
E. Developing Plans for the Use of Equipment, Materials, or Techniques
1. INCLUDED IN PLAN OF INVESTIGATION EQUIPMENT, MATERIAL, OR TECHNIQUES WHICH MET THE REQUIREMENTS OF THE PROBLEM.

This behavior seems to involve:

(1) Identification of the requirements of the problem
(2) The ability to relate pertinent characteristics of equipment, material, or techniques to the problem
(3) The ability to evaluate these characteristics and to select those procedures that will best fulfill the requirements of the problem

An item intended to predict this behavior:

The following item describes a problem, specifying its requirements, and presents sketches of instruments that might be used to fulfill the requirements; the examinee is asked to select the instrument that could fill the requirements most adequately. This should test (2) and (3) above.



Specifications for power transmission wire state that no diameter may be greater than 1.00 inch. The prescribed inspection procedure requires that the diameter of the wire, which is produced in 1000-foot lengths, be checked at one point selected at random in each 10 feet of wire.

In designing an instrument for accomplishing this purpose, which of the basic designs pictured above would provide, with appropriate modifications and refinements, the most adequate instrument?

(Intended Answer: A)

- - - - - -

II. Planning and Designing the Investigation
F. Anticipating Difficulties
5. TOOK SPECIAL PRECAUTIONS IN PLANNING TO PREVENT DAMAGE TO EQUIPMENT.

This behavior seems to involve:

(1) Personal attitude or outside pressure which demands that attempts be made to keep damage to equipment at a minimum
(2) The ability to anticipate where damage might occur in particular equipment
(3) The ability to think of and evaluate techniques that would reduce the possibility of such damage

An item intended to predict this behavior:

Illustrate and describe a piece of equipment and state that it is particularly susceptible to damage. Offer as alternatives various warnings to operators or changes in apparatus and require examinee to select the one most likely to decrease incidence of accidents to apparatus. This should test (3) above.

- - - - - -

II. Planning and Designing the Investigation
G. Determining the Number of Observations
1. OUTLINED PLAN OF INVESTIGATION WHICH PROVIDED FOR SUFFICIENT QUANTITY OF DATA TO BE TAKEN.

This behavior seems to involve:

(1) Concern about whether or not data are sufficient
(2) The ability to judge what would be sufficient or insufficient data for the problem at hand
(3) Including the judgment made in (2) as a part of the plan of investigation

An item intended to predict this behavior:

Describe an experiment in which data insufficient as a basis for conclusions have been collected. Offer five alternatives, four of which indicate that data so far obtained are sufficient or more than sufficient. The fifth alternative should describe a program for the collection of necessary data in the most efficient manner.

- - - - - -

III. Conducting the Investigation
A. Developing Methods, Materials, or Equipment
1. DEVISED AN IMPROVED METHOD, MATERIAL, OR EQUIPMENT.

This behavior seems to involve:

(1) The ability to detect a deficiency in existing methods, materials, or equipment.
(2) A desire to correct the deficiency.

(3) The ability to analyze the source of the deficiency
(4) A store of experience from which to develop possible corrections for the deficiency
(5) The ability to evaluate these corrections in the light of (1) and (3), and to select the most appropriate correction

An item intended to predict this behavior:

Describe a piece of apparatus and the purpose it is to serve. Offer as alternatives changes in the apparatus and require examinee to select the change that would be of greatest value in accomplishing the purpose. This should test (1), (3), and (5).

- - - - -

III. Conducting the Investigation
B. Applying Methods and Techniques
1. USED A TECHNIQUE, MATERIAL, OR EQUIPMENT WHICH SOLVED PROBLEM OR ELIMINATED DIFFICULTY IN THE INVESTIGATION.

This behavior seems to involve:

(1) The ability to think of possible techniques to solve the problem
(2) The ability to predict the consequences of the techniques available
(3) The ability to select the most appropriate technique for the purpose

An item intended to predict this behavior:

Describe a problem concerning the use of a particular instrument. Ask the examinee to select from a list of five techniques the one that would best solve the problem.

- - - - -

III. Conducting the Investigation
C. Modifying Planned Procedures
6. ANTICIPATED THE PROBABLE EFFECT OF A DEFECT OR DIFFICULTY AND TOOK APPROPRIATE ACTION.

This behavior seems to involve:

(1) The ability to foresee possible difficulties
(2) Taking an inventory of available procedures which might remedy the defect
(3) Selecting the remedy that will provide the most satisfactory solution

An item intended to predict this behavior:

Describe a situation in which a defect is developing and offer as alternatives an inventory of available materials or procedures. Examinee must select alternative that will provide the best remedial effect when properly applied.

- - - - -

III. Conducting the Investigation
D. Applying Theory
6. DETERMINED THE PURPOSE OR OPERATION OF A DEVICE BY STUDYING ITS CONSTRUCTION.

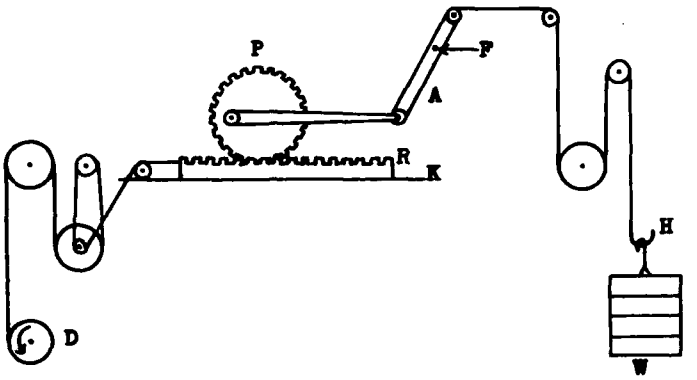
This behavior seems to involve:

(1) Application of knowledge of mechanical principles - for example, relating knowledge of parts of similar devices to new situation and hypothesizing about functions of parts
(2) Testing of hypotheses by physical or mental manipulations of parts
(3) Organizing conclusions about functions of parts to deduce function of apparatus as a whole
(4) Selecting possible applications related to the function and judging practicality of the apparatus for each such purpose.

An item intended to predict this behavior:

The following item presents a diagram of an apparatus and describes the parts; possible changes in one part if another part is manipulated are listed. The exam-

inee is asked to select the particular change that will occur. This should test (1), (2), and (3) above.



Assume that the mechanical device pictured above is friction-free. Let each dot represent a pivot. Assume that the pinion gear, P, is free to rotate about its axle in the center, but otherwise fixed in position. The rack, R, may run along the keyway, K. Point F is a fixed pivot, i.e. the pivot itself cannot change position, although the arm, A, may move about it. The weight, W, is hung on the hook, H, so that it is suspended. The drum, D, has been turned counterclockwise, bringing the device to its pictured state. What would be the effect on the weight if the drum were now released?

A. The weight would fall a short distance and remain stationary.

B. The weight would rise a short distance and remain stationary.

C. The weight would not move.

D. The weight would fall a short distance, then rise.

E. The weight would rise a short distance, then fall.

(Intended Answer: C)

- - - - - -

III. Conducting the Investigation
E. Attending to and Checking Details
3. HANDED IN WORK ONLY AFTER HE HAD CHECKED IT.

This behavior seems to involve:

(1) Concern about accuracy
(2) Willingness to undertake the extra labor involved in checking results

An item intended to predict this behavior:

Present partially completed calculations of experimental results and require examinee to select value to be reported. Calculations so far completed contain an arithmetical error.

- - - - - -

III. Conducting the Investigation
F. ANALYZING THE DATA

This behavior seems to involve:

(1) Information about techniques available and their applicability to various situations
(2) Knowledge of how to apply these techniques

An item intended to predict this behavior:

Present data from an investigation, and offer as alternatives simple ways of analyzing data; have examinee select the most appropriate one.

- - - - - -

IV. Interpreting Research Results
A. Evaluating Findings
2. DREW CONCLUSIONS IN ACCORDANCE WITH CORRECT LOGICAL PRINCIPLES.

This behavior seems to involve:

(1) The ability to understand what is involved in a set of data or propositions
(2) The ability to recognize what is and is not necessarily implied by these data or propositions, or
(3) The ability to formulate a valid generalization from these data or propositions

An item intended to predict this behavior:

The item below presents a series of propositions and asks the examinee to select the two propositions which mean the same thing. The item should sample (1) and (2) above.

1. All forms of pneumonia are caused by viruses.
2. All forms of pneumonia are caused by all viruses.
3. Pneumonia, in all its various forms, is caused only by viruses.
4. Only pneumonia, in all its various forms, is caused only by viruses.
5. Whenever a case of pneumonia occurs, it has been caused by a virus.
6. Diseases other than pneumonia are caused by viruses.

Without reference to the truth or falsity of the statements above, which pair say the same thing?

A. 1, 5
B. 2, 4
C. 3, 5
D. 6, 1
E. 2, 3

(Intended Answer: C)

- - - - - -

IV. Interpreting Research Results
B. Pointing Out Implications of Data
2. WORKED OUT APPLICATIONS TO OTHER PROBLEMS OR FIELDS.

This behavior seems to involve:

(1) Knowledge of the nature of related problems in the same field or in related fields; this includes a general knowledge of the variables, current findings, and theories relative to such problems
(2) The ability to perceive relationships between results of a current problem and related problems

An item intended to predict this behavior:

Present findings that may be expected to have considerable effect in a number of fields. Ask examinee to choose field in which effect will be least.

- - - - - -

V. Preparing Reports
A. DESCRIBING AND ILLUSTRATING WORK

This behavior seems to involve:

(1) The ability to recognize the importance of certain factors in describing and illustrating work, for example, clarity and conciseness

(2) The ability to achieve these factors
(3) The ability to judge when these factors have been achieved

An item intended to predict this behavior:

Describe an experiment and present five alternative ways of stating findings. Only one of these statements should meet standards for conciseness and brevity. Selecting this alternative should demonstrate (1) and (3) above.

- - - - - -

V. Preparing Reports
B. Substantiating Procedures and Findings
4. (Ineffective) OMITTED DETAILS NECESSARY FOR APPLICATION OR CHECKING OF WORK.

This behavior seems to involve:

Recognition of the necessity of including in a report details sufficient for others to apply or to check the work.

An item intended to predict this behavior:

The item below presents a portion of a report that fails to report the methods used in obtaining certain results and lists criticisms that might be applied to the report. The examinee is asked to select the most appropriate criticism. This should sample, in part, the behavior described above.

Recent research in X laboratory has centered around stars whose outermost strata are unusually active. Some of these stars are hot and contain glowing hydrogen in their outer atmospheres. Study of some of these glowing hydrogen atmospheres is particularly difficult since they cannot be seen or photographed. Observations of these atmospheres have led to some calculations of atmospheric motions. The outflow of gas from one star, for example, has been measured as 250 kilometers per second.

Which of the following criticisms is most appropriate for the above paragraph taken from a research report?

A. Inappropriate language is used.

B. Insufficient information is presented.

C. Abstruse theory is presented with no evidence.

D. Excessive detail is included.

E. Correct English usage is violated.

(Intended Answer: B)

- - - - - -

V. Preparing Reports
C. Organizing the Report
7. PLACED LENGTHY OR DETAILED ANALYSIS OR DATA IN APPENDIX.

This behavior seems to involve:

(1) Knowledge of standard practice in organizing research reports
(2) Understanding of the relative advantages of alternative ways of organizing reports

An item intended to predict this behavior:

Present five kinds of material to be included in a report, several of which might be placed in an appendix. The examinee would demonstrate (1) and (2) by selecting the kind of material that most clearly belongs in the appendix.

- - - - - -

V. Preparing Reports
D. Using Appropriate Style in Presenting Report
1. USED A STYLE ADAPTED TO AUDIENCE OR READERS.

This behavior seems to involve:

(1) An appreciation of the styles of composition appropriate to audience
(2) An ability to use the appropriate style

An item intended to predict this behavior:

Present a paragraph presumably taken from a scientific report. The paragraph should contain several inappropriately informal expressions. Ask the examinee to select the most applicable criticism of the paragraph.

- - - - -

VI. Administering Research Projects
A. Selecting and Training Personnel
2. ASSIGNED WORKER A PROJECT SUITING ABILITY OR TRAINING.

This behavior seems to involve:

(1) Obtaining sufficient information about the requirements of the job
(2) Procuring sufficient information about the skills and interests of available personnel
(3) Assigning personnel to specific projects on the basis of information obtained in (1) and (2)

An item intended to predict this behavior:

Describe some qualifications of personnel available for job. Also describe some job requirements. Examinee should base choice on fact that neither type of information is sufficiently detailed for use as a basis for assignment of best worker to the job.

- - - - -

VI. Administering Research Projects
B. Dealing With Subordinates
a. ADMINISTERING REPRIMANDS, RECOGNITION, AND PRAISE.

This behavior seems to involve:

(1) An ability to decide objectively whether reprimand, recognition, or praise is called for
(2) An ability to devise an appropriate reprimand, recognition, or praise
(3) Administering (2)

An item intended to predict this behavior:

Describe a situation in which a subordinate has been guilty of unethical behavior. Offer, as alternatives, methods of dealing with the man. Require examinee to select the alternative that would be most appropriate to the offense and would do most to preserve morale of other workers. This would test (2) above.

- - - - -

VI. Administering Research Projects
C. Planning and Coordinating the Work of Groups
d. UNIFYING RELATED GROUPS.

This behavior seems to involve:

(1) Recognition of the importance of unifying the work of related groups
(2) Information and techniques necessary to effect a functional unification

An item intended to predict this behavior:

The following item sets a problem in which an administrator must select a policy involving techniques to coordinate the work of several divisions in a laboratory. The examinee is given a choice of five policies and is asked to select the one that fills the purpose most adequately.

You have recently been appointed Director of an industrial research laboratory. At the present time it consists of eight divisions with from four to ten professional workers in each division. It is your responsibility to coordinate the work of these divisions, to see that they work together, and to see that each makes an effective contribution to the work of the laboratory as a whole. Which of the following procedures should you use to obtain the most adequate coordination and supervision?

A. Appoint as supervisor the man with the best record in each division who will report directly to you at regular intervals.

B. Select from the laboratory as a whole several men who have the best records and appoint them as supervisors directly responsible to you.

C. Create a separate supervisory division whose members will work in all the divisions at different times and will be directly responsible to you.

D. Designate an administrative supervisor in each division, and arrange to meet regularly with these men as a group.

E. Reorganize the laboratory into twelve divisions so that individual professional workers will receive more adequate supervision.

(Intended Answer: D)

- - - - - -

VI. Administering Research Projects
D. Making Administrative Decisions
b. Following Regulations in Decisions
1. DEVIATED FROM STANDARD ADMINISTRATIVE PROCEDURE TO HANDLE EMERGENCY.

This behavior seems to involve:

(1) Familiarity with standard administrative procedure
(2) The ability to recognize an emergency situation
(3) The ability to make a judgment that the situation requires deviation from standard procedure

An item intended to predict this behavior:

Describe a particular emergency situation in which an administrative decision must be made. Present alternative decisions, one of which calls for deviation from standard procedure to handle the emergency. In selecting this choice, the examinee must perform steps (2) and (3).

- - - - - -

VI. Administering Research Projects
E. Working With Other Groups
2. GOT OPPOSING GROUPS TO AGREE

This behavior seems to involve:

(1) The ability and desire to be tactful
(2) The ability and desire to be persuasive
(3) Fairness

An item intended to predict this behavior:

Describe a situation in which two groups are in disagreement; present alternative courses of action in handling the situation, the best choice giving the most opportunity for (1) and (2) to be exercised.

- - - - - -

VII. Accepting Organizational Responsibility
A. Performing Own Work
b. Accepting Responsibility for Own Work
4. (Ineffective) LEFT INCOMPLETE WORK IN SUCH CONDITION THAT OTHERS COULD NOT COMPLETE IT.

This behavior seems to involve:

(1) The lack of a point of view that holds the goals of the research over other more personal goals and/or
(2) The inability to recognize that leaving work in a condition permitting others to carry it on benefits the over-all goals of the research

An item intended to predict this behavior:

Place an individual in the position of leaving a project and ask him to select the action with highest priority for the limited time remaining, i.e., preparing to leave his work so that successor may continue it most efficiently.

- - - - - -

VII. Accepting Organizational Responsibility
B. Assisting in the Work of Others
6. POINTED OUT INADEQUACIES OF A PLAN, REGULATION, OR POLICY TO GROUP RESPONSIBLE FOR IT.

This behavior seems to involve:

(1) The ability to recognize that a plan, regulation, or policy is inadequate in certain respects
(2) The recognition that something should be done to remedy the inadequacy
(3) The recognition that it is usually most desirable to approach those responsible for the inadequacy first

An item intended to predict this behavior:

The item below describes a situation in which a plan for coordinating the work of two groups is not working efficiently. The work of one group is being delayed by the other group. The examinee is given a choice of several actions that the head of the former group is to take, one of which is to confer with the head of the slow group. This item should sample, in part, (1), (2), and (3) above.

Assume that you are the chief of the Wind-Tunnel Section of a large aeronautical engineering research laboratory. The Test Models Section of the laboratory supplies you with models upon which your section performs its tests. A high priority, rush project has been under way for some time and you have been testing models for three weeks. However, the Test Models Section has been consistently sending you models two or three days later than the schedule agreed upon. This has resulted in considerable loss of time and waste of facilities. Which of the following actions should you take?

A. Explain to the supervisor of the laboratory that your section is not responsible for the delay.

B. Inform the head of the Test Models Section that you intend to report the situation to the supervisor unless it improves.

C. Suggest a readjustment of the schedule to the chief of the Test Models Section so that the two groups can better coordinate their work.

D. Ask the Test Models Section chief to put his section on overtime so that your section can achieve its objectives most effectively.

E. Do nothing since your responsibility extends only to your own section, which is functioning in a way that cannot be criticized.

(Intended Answer: C)

VII. Accepting Organizational Responsibility
C. Subordinating Personal Interests
1. PURSUED MAIN PROBLEM WITHOUT BEING SIDE-TRACKED BY PERSONAL INTERESTS IN OTHER WORK.

This behavior seems to involve:

(1) Holding the goals of a problem over personal interests in other work
(2) The ability to discriminate between behavior directed toward a working goal and behavior performed for personal reasons

An item intended to predict this behavior:

Describe a situation where an individual has become side-tracked from the main goal and is pursuing a personal interest only indirectly related to the goal. In selecting the most appropriate criticism of this behavior, the examinee would demonstrate (2) above.

- - - - -

VII. Accepting Organizational Responsibility
D. Accepting Regulations and Supervision
2. SOLICITED AID FROM SUPERIOR WHEN FACED WITH DIFFICULTIES.

This behavior seems to involve:

(1) Recognizing a situation that individual is not capable of coping with alone
(2) Realizing that it is most effective to request aid immediately without wasting time

An item intended to predict this behavior:

Describe a situation calling for a decision that an individual is not capable of making. From five possible lines of action, the examinee must select the one that suggests asking for help from a superior. In doing so, the examinee must perform (1) and (2) above.

- - - - -

VIII. Accepting Personal Responsibility
A. Adapting to Associates
1. CRITICIZED THE WORK OF ANOTHER WITHOUT DISPARAGING HIS EFFORTS.

This behavior seems to involve:

(1) A recognition that the feelings (egos) of others are involved in their work, and that individuals perform better if they are not <u>personally</u> criticized
(2) The ability to offer a criticism or suggestion tactfully through such techniques as (a) first discussing the problem to learn the individual's general approach, knowledge, and opinions about the problem, and (b) bringing in the criticism or suggestion at an appropriate time and in a tactful manner

An item intended to predict this behavior:

Present a situation in which the work done by one staff member has definite shortcomings. Allow examinee to decide whether or not he, as a co-worker, would criticize the work of the other, and if so, whether he would select the choice describing a criticism least calculated to belittle the other.

- - - - -

VIII. Accepting Personal Responsibility
B. Adapting to Job Demands
2. CONTINUED TO WORK ACCORDING TO PLAN IN SPITE OF OBSTACLES OR OPPOSITION.

This behavior seems to involve:

(1) A continuing strong conviction that the plan of a particular piece of work is the best available for that work.
(2) The skills necessary to act on that conviction in the face of opposition

(3) The ability to avoid being dissuaded from fulfilling that conviction

An item intended to predict this behavior:

Describe a situation in which a man is convinced that his present way of attacking a problem is the plan that should be followed; also describe various obstacles to fulfilling the plan. The examinee must choose to continue with the plan when offered five choices of action to be taken.

- - - - - -

VIII. Accepting Personal Responsibility
C. Meeting Personal Commitments
1. (Ineffective) FAILED TO PERFORM A SPECIFICALLY ASSIGNED TASK TO WHICH HE HAD AGREED.

This behavior seems to involve:

(1) Concern for meeting commitments
(2) Willingness to put forth considerable effort in order to meet commitments

An item intended to predict this behavior:

The item below describes a situation in which an individual has made a commitment but finds that he will have considerable difficulty in meeting it. A list of alternative lines of action is presented. The examinee should choose the alternative indicating that the individual will meet his commitment even though it involves personal inconvenience. In order to prevent the item from being too obvious the correct answer is limited to only a part of what is considered to be the best action in this situation.

While a research group is developing an improved steering mechanism for guided missiles, the director is requested to attempt modification of a particular unit in the device so that it could also be used for certain other purposes. First priority is to be given to the main project, but it is considered "highly desirable" that the sub-project be carried out concurrently. Since the main schedule is very tight, the director asks for a volunteer from the staff--someone who is temporarily ahead of schedule on his phase or who can manage to take on the additional assignment by making special arrangements. Dr. North is the only volunteer, and the sub-project is turned over to him. Shortly thereafter, Dr. North finds that unforeseen changes have to be made in the schedule for his phase of the main project. It now appears that the time he had planned to spend during the next few weeks on the sub-project will have to be devoted to the main project.

If Dr. North were limited to the alternatives listed below, which of these actions should he take?

A. Complete the sub-project, working evenings and week-ends if necessary.

B. Inform the director that it will be necessary to get someone else to take over the sub-project.

C. Go to another staff member and ask him to take over the sub-project.

D. Shift his main efforts to the sub-project temporarily, and finish it before continuing work on the main project.

E. Recommend to the director that work on the sub-project be delayed until successful completion of the main project is assured.

(Intended Answer: A)

- - - - - -

VIII. Accepting Personal Responsibility
D. BEING FAIR AND ETHICAL

This behavior seems to involve:

(1) Considering or presenting <u>all</u> evidence objectively

(2) Being truthful in all communications
(3) Giving credit where credit is due
(4) Taking responsibility for own actions

An item intended to predict this behavior:

Describe a situation where criticism may be expected if (1) is done; examinee must recognize that (1) is nevertheless necessary. This recognition of the necessary action does not, of course, assure that the examinee would actually perform comparable behavior.

- - - - -

VIII. Accepting Personal Responsibility
E. Showing Interest in Work
1. WORKED ON OWN TIME TO IMPROVE TECHNICAL KNOWLEDGE AND PROFICIENCY.

This behavior seems to involve:

A strong interest in and desire to learn more about one's work

An item intended to predict this behavior:

Describe briefly the work of a group of five research workers concerned with the same project. Include in the description the actions of each which might indicate interest in the work, the action of one being extensive study of material related to the problem on his own time. The actions of the others should be less outstanding. Ask the examinee to evaluate the men by selecting the one who the evidence suggests would be the most desirable worker. The item would measure recognition of the value of personal effort to improve own technical knowledge and proficiency.

CHAPTER V

C O N C L U S I O N

An account of the objectives, planning, development, and results of the test development project has been given in preceding chapters. The work was carried out as part of a broad research program which would include:

1. Determination of the critical requirements for successful participation in research and engineering work.
2. Development of an aptitude test for the selection of scientific personnel.
3. Development of tests to measure proficiency in specific areas of scientific work.
4. Development of procedures for evaluating the job performances of scientific personnel.
5. Determination of the predictive value of the tests developed in steps (2) and (3) using the procedures developed in step (4) to obtain evaluations of personnel for comparison with test predictions.

The first of these steps was completed in an earlier project, while the second has been the subject of this report. Objectives of the present project were, first, that a test be developed which would be applicable to the selection of candidates for advanced training in natural science and engineering, as well as to the selection of junior professional workers in research laboratories; second, that it be a test of potentiality for, rather than proficiency in, research work; and third, that each test item developed be an attempt to predict a specific critical behavior identified in a preceding study concerning the critical requirements for research personnel.

A number of working assumptions were made explicit in planning the test:

1. That the critical behaviors identified in the preceding study represent the most important job behaviors of scientific personnel and should therefore serve as the basis for building test items.
2. That most of the important job behaviors are common to all the subject-matter areas found in research and engineering work, and that most of the aptitudes underlying these behaviors are also common to the various scientific fields.
3. That each test item should grow directly from a specific critical behavior by posing a problem situation requiring insofar as possible that particular behavior to be demonstrated.
4. That examinees would have a minimum common background of one year's college training in physics, chemistry, and mathematics; all necessary technical information beyond this level would be given in the statement of each problem; the total test would be balanced in subject-matter content so that no one field of specialization would be involved in a majority of items.
5. That the multiple-choice item with five alternatives would best suit project purposes; however, all varieties of objective test items were considered throughout the item-development phases.

As a preliminary to construction of test items, each critical behavior was studied, and ideas for appropriate types of test items to predict each behavior were formulated. Current objective tests designed to measure some aspect of scientific aptitude were also reviewed. The relative proportion of critical behaviors reported for each major area in the preceding study was used as a rough guide to the number of items required for each category.

Items were prepared by individuals with knowledge of both scientific subject-matter and the technical aspects of test-construction. During the editing process the following criteria were used:

1. Is it probable that the item will predict success or failure on the job as defined by the critical behaviors?
2. Is the difficulty-level of the item appropriate to the group for which the test is intended?
3. Is the answer intended as correct clearly the best choice of those offered?
4. Does solution of the item require only such academic information as is ordinarily acquired in one-year college courses in physics, chemistry, and mathematics?
5. Does the item meet the generally accepted technical requirements of good test items?

After preliminary editing, all items were reviewed independently by several persons professionally engaged in some area of natural science or engineering. A review of final tryout forms of the test was made by four editorial consultants. Trial administrations with groups of examinees similar to those for which the test is intended were then conducted. Distributions of scores and difficulty-levels of items were presented. The results of these trial administrations were used as the basis for further revision of items.

The final form of the test consists of approximately 170 items organized in three groups or subtests. Subtest I is intended to predict chiefly job performances related to formulating problems and hypotheses and planning and designing the investigation; subtest II is intended to predict behaviors related to conducting the investigation and interpreting research results; and subtest III, behaviors related to preparing reports, administering research projects, and accepting organizational and personal responsibility.

For security reasons the test items are not contained in this report. However, sample items are presented which illustrate the types of items and subject-matter content contained in the test proper. A larger number of "rationales" or descriptions of item-types are also presented; these contain a formulation of hypotheses as to the knowledges, abilities, or personality traits believed to underlie a specific critical behavior, and a description of a type of test item believed to predict that behavior. These rationales were also reviewed by editorial consultants.

It is believed that the test will be of value in selecting research personnel for the following reasons:

1. Test items were developed to measure specific behaviors identified by scientists themselves in a preceding study as crucial to successful performance in research work.
2. Items were prepared by individuals with knowledge of both scientific subject-matter and the technical aspects of test-construction.
3. Items were criticized by persons professionally employed in research work in natural science and engineering, and were revised on the basis of such review.
4. Trial administrations were conducted to determine the difficulty of items and to identify further editorial needs.
5. Items were reviewed by editorial consultants with particular reference to whether each would predict the intended critical behavior.

The test of research aptitude developed in this project has been reproduced and is available for use in the later phases of the long-range research program which has been outlined.

B I B L I O G R A P H Y

1. Benton, A. L. and Perry, J. D. A Study of the Predictive Value of the Stanford Scientific Aptitude Test. Journal of Psychology, 1940, 10, pp. 309-312.

2. Berdie, Ralph F. The Differential Aptitude Tests as Predictors in Engineering Training. American Psychologist, 1949, 4, p. 292.

3. Bernreuter, R. G. and Goodman, C. H. A Study of the Thurstone Primary Mental Abilities Tests Applied to Freshmen Engineering Students. Journal of Educational Psychology, 1941, 32, pp. 55-60.

4. Burke, Paul J. Testing for Critical Thinking in Physics, Unpublished paper, January 1949.

5. Castore, George. A Screening and Selection Battery for Prospective Physicists and Chemical Engineers. Unpublished thesis, The Pennsylvania State College, 1948.

6. Chrisof, C. The Formulation and Elaboration of Thought-Problems. American Journal of Psychology, 1939, 52, pp. 161-185.

7. Crawford, A. B. and Burnham, P. S. Forecasting College Achievement. New Haven: Yale University Press, 1946.

8. Davis, Frederick B. (Ed.) The AAF Qualifying Examination. Army Air Forces Aviation Psychology Program Research Reports, No. 6, Washington: Government Printing Office, 1947.

9. Davis, J. C. The Measurement of Scientific Attitudes. Science Education, 1935, 19, pp. 117-122.

10. Downing, E. R. Some Results of a Test of Scientific Thinking. Science Education, 1936, 20, pp. 121-128.

11. Englehart, Max D. and Lewis, Hugh B. An Attempt to Measure Scientific Thinking. Educational and Psychological Measurement, July 1941, 1, No.3, pp.289-294.

12. Flanagan, John C. et al. Critical Requirements for Research Personnel. Pittsburgh: The American Institute for Research, 1949.

13. Gamble, Allen O. The Nature of Research Work and of Research Workers in the Physical Sciences. Paper presented before the Annual Meeting of the American Psychological Association, September 1949.

14. Glaser, E. M. An Experiment in the Development of Critical Thinking. Teachers College Contributions to Education, No. 843. New York: Columbia University Press, 1941.

15. Griffin, C. H. and Barow, H. An Engineering and Physical Science Aptitude Test. Journal of Applied Psychology, October 1944, 28, pp. 376-387.

16. Hamor, William A. Human Aspects of Scientific Research. Science, September 1945, 102, No. 2645, pp. 237-241.

17. Howard, Frederick T. Complexity of Mental Processes in Science Testing. Teachers College Contributions to Education, No. 879. New York: Columbia University Press, 1943.

18. Kilpore, William Arlow. Identification of Ability to Apply Principles of Physics. Teachers College Contributions to Education, No. 840. New York: Columbia University Press, 1941.

19. Laycock, S. R. and Hutcheon, N. B. A Preliminary Investigation into the Problem of Measuring Engineering Aptitude. Journal of Educational Psychology, 1939, 30, pp.280-288.

20. Mandell, Milton M. The Selection of Foremen. Educational and Psychological Measurement, 1947, 7, pp. 385-397.

21. Mercer, M. An Analysis of the Factors of Scientific Aptitude as Indicated by Success in Engineering Curricula. Pennsylvania State College Studies on Education, 1940, 22, pp. 47-48.

22. Moore, Joseph E. A Decade of Attempts to Predict Scholastic Success in Engineering Schools. Occupations, November 1949, 28, No. 2, pp. 92-96.

23. Smith, Eugene R. Appraising and Recording Student Progress, Vol. III of "Adventures in American Education." New York: Harper and Brothers, 1942.

24. Vaughn, K. W. The Graduate Record Examinations. Educational and Psychological Measurement, 1947, 7, pp. 745-746.